

**U.S. PATENT APPLICATION**

**for**

**COMPUTER-AIDED OPTIMIZATION OF SUBSTANCE LIBRARIES**

**Inventors:**      Dr. Stephan A. SCHUNK  
                            Dr. John M. NEWSAM

## COMPUTER-AIDED OPTIMIZATION OF SUBSTANCE LIBRARIES

### BACKGROUND OF THE INVENTION

#### 1) Field of the Invention

**[0001]** The present invention relates to an apparatus and a method for optimizing substance libraries comprising at least two substances.

**[0002]** More precisely, the invention relates to a computer-aided method which permits the successive material optimization of non-molecular substance libraries with respect to a defined application and/or the optimization of the test parameters which are relevant to the application or are important to the process of selected substances within a predefined parameter space or a parameter space resulting during the optimization (altered with respect to the predefined parameter space), an apparatus suitable therefor, comprising a substrate, the substance library per se which can be produced in this way and a computer program for controlling the inventive method.

**[0003]** In the context of the inventive method it is possible to optimize substance libraries comprising molecular and non-molecular substances.

#### 2) Description of Related Art

**[0004]** The preparation of substance libraries in pharmaceutical combinatorial research generally begins with designing a hypothesis regarding the interaction of an active compound molecule (for example a ligand) with a biological target (receptor). The type and strength of interaction here are associated with discrete structural properties of the active compound molecule (SAR – structure activity relationship). The term "structural property" includes here, for example, topology, conformation, spatial arrangement of substituents or electronic configuration of the active compound molecules. These "structural properties" are incorporated as descriptors, that is to say as

parameters for describing the interaction(s) between active compound molecule and receptor. The classical combinatorial approach is based on the systematic change (permutation) of the structural properties of one or more (molecular) organic base skeletons. This presupposes the existence of suitable organic building blocks (for example a methyl or phenyl group) which react under defined synthesis conditions with the base skeleton to form a target molecule of the planned substance library. After determining the activity of the molecules synthesized within the substance library against a biological target, the hypothesis originally made is revised with the objective of producing an optimized substance library.

[0005] US 5,901,069 and 5,463,564 disclose systems and methods for the at least partial automated generation of (chemical) compounds having desired chemical or bioactive properties.

[0006] There, starting from an initial hypothesis relating to the interactions of interest and the structural features required for this, a computer-aided process is carried out which during each iteration comprises the following steps:

[0007] (1) a library consisting of a plurality of compounds is robotically generated in accordance with robotic system instructions;

[0008] (2) the compounds in the library are analysed in order to identify those compounds which have the desired useful properties;

[0009] (3) structure-activity data are utilized in order to select compounds which are to be synthesized in the next iteration; and

[0010] (4) new robotic system instructions are generated by the experimenter which control the synthesis of the compounds in the library for the next iteration.

[0011] As an aid to refining the hypothesis initially made in pharmaceutical research, suitable software is available for modeling and visualizing molecules, or else mathematical/statistical software which comprises, for example, regression methods, for example linear single or multiple regression, for quantifying the structure-activity relationship (QSAR – quantitative structure

activity relationships). In the literature, corresponding computer-aided methods for pharmaceutical applications are also termed CADD (computer-aided drug design) or CAMD (computer-aided molecular design).

[0012] The systems and methods disclosed in these publications are not fully automated for optimization, that is to say the experimenter must intervene at one or more time points when the method is being carried out. In addition, the change(s) made to the "structural property" in these methods is (are) always carried out by varying discrete states, that is to say for example dependent on the variation of a substitution pattern, for example methyl → ethyl → propyl → ...

[0013] In summary, the above described implies that the targeted optimization, which is not based exclusively on trial and error, of molecular organic libraries is based on the following principles:

[0014] a) assuming a relationship between structure and activity of a molecule in question (SAR);

[0015] b) existence of suitable synthetic building blocks;

[0016] c) variation of discrete molecular properties;

[0017] d) use of molecular descriptors.

[0018] Whereas points a) and d) do not represent a fundamental precondition for the combinatorial variation of an organic base skeleton, points b) and c) are an essential necessity therefor.

[0019] Since in the case of non-molecular substances the required relationships between structure and activity of the substances are often unknown, and corresponding molecular descriptors do not exist, for the complete integration of the generation of non-molecular substance libraries, recourse cannot be made to the methods for library optimization in pharmaceutical active compound research.

[0020] The preparation and testing of non-molecular substance libraries outside pharmaceutical research is described in a range of publications. In this sector, production, testing and evaluation of the substance libraries are represented as separate process steps. However, the complete integration of

the individual process steps in to a joint software environment has not yet been described to date, as can be seen from the following summary of the relevant prior art:

[0021] Danielson et al. describe a combinatorial method for discovering and optimizing luminescent substances (Nature, vol. 389, p. 944, 1997). The method of Danielson comprises the automated production of a first substance library, testing the first substance library in order to identify lead materials, and designing and synthesizing newly optimized substance libraries by the experimenter on the basis of the composition of the lead substances identified. According to Danielson, the combinatorial exploration requires different iterations in order to optimize the composition and production for a defined application. The iterative optimization according to Danielson is based, however, solely on the intuition of the experimenter without the use of computers or the use of software-controlled optimization methods.

[0022] WO 00/23921 relates to a computer-controlled method for generating a library design for a combinatorial material library which comprises:

[0023] defining one or more sources and one or more destinations, each source being an electronic data point representing a component for preparing the combinatorial library and each destination being an electronic data point representing an arrangement of cells;

[0024] receiving an input which defines the first mapping, this first mapping being electronic data defining a distribution pattern for assigning a component to cells in the arrangement, the distribution pattern defining a minimum and a maximum amount of the component and a gradient between the minimum and the maximum amounts of the component across the multiplicity of the cells;

[0025] using the first mapping to calculate a composition of one or more materials to be assigned to one or more of the cells; and

[0026] generating a data file for defining the library design, the data file comprising electronic data representing the sources, the destinations and the mapping.

[0027] According to this publication, therefore, only the automated production of material libraries is described or claimed.

#### SUMMARY OF THE INVENTION

[0028] One object of the present invention is to provide an integrated system and method for producing and optimizing non-molecular substance libraries which makes it possible to optimize a substance library without intervention, or with only marginal intervention, for example in the form of a plausibility analysis of an operator.

[0029] This object and other objects are achieved according to the invention by a method for the automated production and iterative automated optimization of a substance library and/or at least one reaction parameter comprising at least two substances comprising the steps:

[0030] a) defining at least one production parameter and at least one test parameter;

[0031] b) automated preparation of a substance library by producing at least two substances on the basis of the at least one production parameter;

[0032] c) automated testing of the at least two substances of the substance library with respect to at least one wanted useful property on the basis of the at least one test parameter;

[0033] d) evaluating the test using electronic data analysis;

[0034] e) varying the at least one production parameter and/or the at least one test parameter for optimizing the wanted useful properties, and single or repeated repetition of steps b) to e) or c) to e);

[0035] wherein the steps b) to e) are carried out in an integrated automated process and by

[0036] an apparatus for the automated production and iterative automated optimization of a substance library, comprising:

[0037] a) means for defining at least one initial production parameter and at least one test parameter;

[0038] b) an arrangement for the automated production of the at least two substances of a substance library on the basis of the at least one production parameter;

[0039] c) a test device for the automated testing of the at least two substances of the substance library with respect to at least one performance characteristic on the basis of the at least one test parameter;

[0040] d) a data analysis system for evaluating the test;

[0041] e) device for varying the at least one production parameter and/or test parameter for optimizing the performance characteristics, characterized in that

[0042] f) the apparatus in addition comprises means for the integrated and iterative automated control of steps b) to e) or c) to e).

[0043] Further objects, features and advantages of the present invention, will become readily apparent from detailed consideration of the preferred embodiments which follow.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0044] Fig. 1 shows a flow diagram of an embodiment of the inventive method;

[0045] Fig. 2 shows the relationship between production parameters and test parameters and the performance characteristics of a catalyst;

[0046] Fig. 3 summarizes diagrammatically the possible interventions by an operator of the inventive apparatus;

[0047] Fig. 4 shows the data flow and material flow when the inventive method is carried out;

[0048] Fig. 5a shows the design and test of an initial substance library;

[0049] Fig. 5b shows an analysis method for restricting the parameter space;

[0050] Fig. 5c shows the principle of an evolutionary algorithm;

[0051] Fig. 6 summarizes the sensitivity analysis of the example of Table 2;

[0052] Fig. 7 shows the sensitivity analysis for the first optimized substance library from Table 3.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0053] The basis for the inventive method and the inventive apparatus is the correlation of sets of at least one production parameter and/or test parameter ("experimental descriptors") and optionally characteristic properties ("first or second order characteristics") of the substances of the substance libraries with their performance characteristics within a method integrated with respect to control software, data capture software and data evaluation software and a corresponding apparatus.

[0054] The optimization of the non-molecular substance libraries and/or reaction conditions is carried out in an iterative process in which an initially given non-molecular substance library which is characterized by production properties and thus characteristic properties, comprising at least two substances, is initially subjected to one or a multiplicity of precisely defined tests of defined performance characteristics. By applying mathematical algorithms, the significance of the individual parameters in the production and/or test parameter sets and/or characteristic properties to the test result is tested, the appropriate parameter space is constricted or changed and optimized parameter sets are prepared. The optimized (generally changed with respect to the predefined parameter sets) parameter sets then form the basis for the preparation of the production parameter or production parameters for the substances of the next substance library and/or for establishing new test parameters. To obtain such optimized parameter sets, information from one or more libraries can be used or else evolutionary algorithms and other mathematical or statistical methods can be used.

[0055] In addition, in the context of the inventive method, not only different substance libraries can be studied and optimized under the same test conditions, but the same substance libraries can be studied and optimized under different test conditions.

[0056] Furthermore, the interaction between the (chemical) constitution of the substances within the optimized substance libraries and the possible variation of the test parameters, there also results the possibility of simultaneous optimization of the substances of the material library with respect to their performance characteristics and test conditions which typically relate to the optimization of at least one reaction parameter with respect to later use of one or more substances within the substance library, in which case here likewise the work can be carried out in such a manner that not only are the substances within the substance library and the at least one reaction parameter optimized simultaneously, but, with respect to a defined, typically previously optimized, substance library, further attempts are made to optimize the reaction parameters.

[0057] The iterations can be repeated until either the explorative destination is reached or the optima within one or more preset parameter space or parameter spaces are reached. The efficiency of the inventive method COALA (Computation Assisted Library Amendment) results from the integration of the production parameters and test parameters and of the characteristic properties and test data in a shared software environment preferably under real time conditions.

[0058] It is a further advantage of the inventive method that this can be used for an optimization over short and also long periods, that is to say either a substance library is optimized (short period) or optimization is performed over the total of a plurality of substance libraries (long period). A parameter set for an improved library is always obtained after processing of a substance library.

[0059] A diagrammatic description of one embodiment of the inventive method is given in Fig. 1. In the figure

- [0060] (1) designates an initial substance library,
- [0061] (2) designates the automated production,
- [0062] (3) designates the automated testing,
- [0063] (4) designates data capture,
- [0064] (5) designates data analysis,
- [0065] (6) designates amended substance library and
- [0066] (7) designates determination of the characteristic properties of the substances within the substance library.

[0067] The terms used in the context of the present application may be clarified at this point:

[0068] Substance library: The term "substance library" designates an arrangement comprising at least two, preferably up to 10, further preferably up to 100, in particular up to 1000, and further preferably up to 100,000 substances, or (chemical) compounds, mixtures of (chemical) compounds, materials, formulations which are present in the solid, liquid or gaseous state on/in a substrate.

[0069] Preferably, in the context of the present invention, substances of the above meaning are non-gaseous substances, for example solids, liquids, sols, gels, waxy substances or substance mixtures, dispersions, emulsions, suspensions and solids, particularly preferably solids. In the context of the substances used according to the invention, the substances can be molecular and non-molecular chemical compounds or formulations or mixtures or materials, the term "non-molecular" defining substances which can be continuously optimized or amended, in contrast to "molecular" substances whose structural property may only be changed by varying discrete states, that is to say for example varying a substitution pattern.

[0070] The substances in the substance library can be identical or different from one another, the latter being preferred; however, in an optimization of test parameters or reaction parameters or process parameters, it is also easily possible that the substance library comprises two or more identical substances or consists exclusively of identical substances.

[0071] Virtual substance library: This is a substance library which has been designed by or using a computer and in which it still remains to be tested, for example by an operator or computer, as to whether the substances or substance classes within the substance library can actually be prepared. In the context of a virtual library, on the basis of available information with respect to the requirements of a substance when used in a defined chemical, physicochemical or physical reaction, for example, a substance library can be one consisting of substances which have been "freely" generated by the computer on the basis of these data.

[0072] Initial substance library: The substance library designed at the beginning of the inventive method for optimization; this can correspond to the virtual substance library, can be modified in comparison with this or be designed by an operator without including the preparation of a virtual substance library.

[0073] Amended substance library: Substance library improved by establishing new production and/ or test parameters which is then subjected to an inventive iteration (optimization).

[0074] Substance: Unit situated in the respective substrate sections within the substance library, and which can consist of one or more components.

[0075] Substrate: Device having a rigid or semi-rigid surface which can be either flat or also have recesses or bore holes or channels. The substrate must be suitable for separating the at least two substances physically from one another in the at least two different sections separated from one another. The substances can be arranged one-dimensionally, two-dimensionally or three-dimensionally in the substrate, that is to say next to one another and one above the other in different planes.

[0076] Preferably, the substrate comprises parallel continuous channels and can have, inter alia, a wire grid or a foamed ceramic.

[0077] Further preferably, the substrate is a tube-bundle reactor, in particular a tube-bundle reactor as described in WO 99/41005 and its

disclosure in this respect is incorporated by reference in its entirety into the context of the present application.

[0078] Performance characteristics: These are measurable characteristics of the substances of the substance library which can be determined within an automated testing and which are optimized in the context of the inventive method: examples of these are mentioned in the further course of the description.

[0079] Automated production: Preparation of the substance library and of the substances of the substance library using substantially automated preparation modules which are incorporated in a suitable control software.

[0080] Automated optimization: Carrying out one or more automated optimization steps.

[0081] Test parameters: One or more changeable variables which establish the type and conditions of the automated testing.

[0082] Reaction parameters: Special case of the test parameter which, if it is in the case of the test the performance of the substance library in the context of a defined (chemical) reaction, establishes the type and conditions of this reaction.

[0083] Analysis parameters: Further subgroup of the test parameters which establishes the type and conditions in the determination of the characteristic properties of the substances within the substance library.

[0084] Parameter space: Totality of all parameters which describe the production and/or testing of the substance library within the context of individual steps or the totality of the inventive method, with in this case only preferably the strictly mathematical or scalar definition of a collection of non-redundant vectors being used for the description of the test parameters and/or reaction parameters and/or production parameters, that is to say the parameter space can also comprise redundant vectors or scalars.

[0085] Output parameters: Analogue or analogue to digital converted data from the production, test and characterization instruments to the file server.

[0086] Analysis: Term which comprises all of the analytical methods for testing substances within a substance library for determining their characteristic properties.

[0087] Characteristic properties: This term represents the structural features or the description of the structural features of a non-molecular material by employing physicochemical methods. "First order characteristics" are very largely those characteristic properties which are obtained using physical characterization methods, for example X-ray diffraction, LEED structure determination, EDX, X-ray fluorescence analysis; X-ray photoelectron spectroscopy, auger spectroscopy.

[0088] "Second order characteristic" is taken to mean those characteristic properties which are accessible using physicochemical characterization methods, for example nitrogen adsorption – (surface dimensions, (BET)); TPD – (bonding strengths of adsorbates to surfaces or selective chemisorption – size of the surfaces of active centres).

[0089] Automated testing: Collective term for tests of one or more performance characteristics which proceed automatically. Automated testing is completely integrated into the software.

[0090] Data acquisition: Capture of all parameters (production or test parameters), characteristic properties and test results.

[0091] Test results: Data sets such as production and/ or test parameters, characteristic properties and performance characteristics, in particular with respect to the characteristics to be optimized.

[0092] Data analysis: Computer-aided data evaluation using various mathematical/ statistical or evolutionary algorithms.

[0093] Sensitivity analysis: Mathematical method for differentiating statistically significant factors from non-significant factors within linear or non-linear equation systems.

[0094] Virtual screening: Optimization of substance libraries using data processing systems using databases and modeling by means of software - supported analytical methods.

[0095] Modified production parameters and/or test parameters:

Production parameters and/or test parameters which have been modified in comparison with the initial production parameters and/or test parameters and have generally been optimized, for example with process optimization as a target.

[0096] Software: Software which comprises the components for process control, data transfer, data formatting, data storage, data search, data management and for data interpretation and data evolution and is preferably operated under real-time conditions.

[0097] Input parameters: Digital or digital to analogue converted data from the control computer/process control system for controlling the production, test and characterization instruments (information).

[0098] Output information: Digital or digital to analogue converted data which corresponds to the results of the testing of the substance library in the respective optimization step.

[0099] Optimization step: Sequence of production and testing the substance library including analysis of the test results which is incorporated in a suitable control and evaluation software.

[0100] Production module: Totality (arrangement) for the automated production of the substances within the substance library.

[0101] Test module: Device or arrangement for the automated testing of the performance characteristics of the substances and/or optimization of one or more reaction parameters.

[0102] System time/date: The time coordinate which is preset by the data processing system and makes possible unambiguous alphanumeric assignment of all processes occurring during the method.

[0103] System location: Location which is unambiguously defined within the totality of the arrangement for production, testing and analysis of the substances.

[0104] The present invention enables a) the information (input parameter and output information) which is obtained in the process steps shown in Fig. 1

(i.e. automated synthesis, analysis and automated testing) to be integrated in a collective software environment and b) in the sense of a self-optimizing method, using mathematical algorithms on the output data, to arrive at a successive improvement in the production parameters and/or test parameters of the subsequent (=optimized) non-molecular substance libraries. In principle the inventive method also offers the possibility of using the characteristic properties of the substances of the substance library ("1<sup>st</sup> and 2<sup>nd</sup> order") as a basis for the optimization of the substance library. The latter embodiment is difficult to achieve especially in the area of heterogeneous catalysis:

[0105] a) Characteristic properties of non-molecular catalysts are always to be taken as an average over a virtually infinitely large structure and do not characterize the individual active centre of an atomic level. Since the totality of the active centres determines the activity and selectivity of a catalyst, an average, in contrast, may not be usable for the desired correlation.

[0106] b) It is always questionable whether all of the characteristic properties which determine the activity and selectivity of a catalyst can be determined. In every case, the time consumed to carry out corresponding measurements and for modelling based on this would be considerable.

[0107] c) If the knowledge of the optimum qualities of the active centres were actually available, it would nevertheless be questionable as to whether a corresponding material could actually be synthesized by suitable choice of the production parameters.

[0108] d) The characteristic properties of a material can differ greatly from one another during the analysis and testing owing to shaping and conditioning processes.

[0109] For these reasons, in the present invention, it is not primarily the characteristic properties ("first and second order characteristics") which are used as a basis for library optimization but preferably the production parameters ("experimental descriptors"), since these, regardless of the structural qualities of the substances, permit correlation with the performance characteristics. Procedures as used in pharmaceutical active compound

research for optimizing molecular structures, that is to say based on the characteristic parameters ("first and second order characteristics"), can, for the reasons a) to c) given above, only be used in exceptional cases for optimizing non-molecular substances or formulations or materials. Examples within non-molecular material research in which the characteristic properties ("first and second order characteristics") can at least partially be correlated with the performance characteristics are:

[0110] a) zeolites: acidity and pore diameter correlate with cracking activity and/or isomerization activity and selectivity;

[0111] b) structure-insensitive reactions: Fischer-Tropsch activity and selectivity is frequently correlated with the size of the material particle situated on the catalyst surface;

[0112] c) semiconductors: UV-spectra provide conclusions with respect to electronic transitions;

[0113] d) pigments: layer thicknesses can correlate with the colour.

[0114] Examples a) to d) represent only a very small section within the bandwidth of non-molecular substances. The relationship between production parameters and test parameters and characteristic properties and the performance characteristics of a catalyst is shown in Fig. 2.

[0115] As can be seen from this figure, initially the synthesis (production parameters) is directly responsible for the first order catalyst characteristics, for example crystallinity, defect structure or particle size distribution. These first order catalyst characteristics then in turn determine the second order catalyst characteristics, for example redox characteristics, BET surface area, heat of adsorption or acidity. Both the first order catalyst characteristics and the second order catalyst characteristics then determine the performance characteristics which obviously in turn depend on the other experimental descriptors, that is to say the reaction parameters, for example reactor and the test parameters, for example reaction conditions and starting material composition.

[0116] The present invention indicates a route which permits

[0117] 1. logical links to be made between the production parameters and/or if appropriate also characteristic properties of substances of non-molecular, non-trivial substance libraries and their performance characteristics;

[0118] 2. logical links to be made between the test parameters of substances of non-molecular, non-trivial substance libraries in a multidimensional test parameter space and the performance characteristics of these substances,

[0119] where the software environment, with only a small amount of intervention by the user is able to repeat the process steps production, testing and if appropriate analysis each time with an improved or mutated set of production parameters or test parameters until the optimum of a material or the optimum of the test conditions has been experimentally found or can be mathematically predicted. The role of the operator (Fig. 3) is essentially restricted to defining chemically useful initial libraries, to operating manually diverse production modules whose full automation is technically difficult or is currently impossible to achieve, and to testing the results of the computer-aided data evaluation for plausibility. With advancing automation, it is absolutely conceivable to carry out the inventive method in a fully automated manner, that is to say the operator inputs the desired performance characteristics and test parameters, while the inventive method carries out optimization of the substance library starting from a virtual substance library. The operator then receives the results of the data analysis and can thus obtain the optimized substance and the optimum reaction conditions and can then immediately implement this knowledge, for example, on a laboratory or pilot-plant scale. Although this fully automated procedure is already possible even now, it fails in many cases due to the fact that the data required to prepare a virtual substance library are not available or are insufficient and knowledge on the relationship between characteristic properties and performance characteristics is insufficient or unavailable.

[0120] Preferably, as shown in Fig. 4, the characteristic properties are captured in parallel to the production parameters and test parameters and

deposited in a shared database. By chemical plausibility testing, then, if appropriate, the suitability of characteristic properties for optimizing non-molecular libraries must be decided on a case-by-case basis.

[0121] The methodology illustrated in Figs. 1 and 4 may be described as follows:

**[0122] Definition of the initial substance library:**

[0123] This requires establishment of the production parameters by which the production of the substances within the substance library is unambiguously describable. The production parameters of the initial library can differ greatly, for example on a logical basis, that is to say based on parameters taken from the literature, on production parameters considered expedient chemically or else on the basis of mathematically determinable production parameters for non-molecular substances and are shown by way of example in Table 1 for three different production methods (that is to say 1. precipitation, 2. metal adsorption and 3. incipient wetness). Each individual production parameter can affect the characteristic properties and thus indirectly also the performance characteristics of the substance. Apart from a few exceptions, for example the type of precursor or support material, every other production parameter can be varied semicontinuously. Here there is a marked difference to the optimization of organic molecules discussed at the outset which is based on the variation of discrete molecular structures.

[0124] The substances within a substance library used according to the invention are not restricted in a particular manner. The substances can be molecular, non-molecular or else formulations and materials. Those which may be preferably mentioned are the following classes of such substances:

[0125] heterogeneous or heterogenized catalysts, luminophores, electrooptical, superconducting or magnetic substances, or mixtures of two or more thereof; in particular intermetallic compounds, oxides, oxide mixtures, mixed oxides, ionic or covalent compounds of metals and/or non-metals, metal alloys, ceramics, organometallic compounds and composite materials, dielectrics, thermoelectrics, magnetoresistive and magnetooptic materials,

organic compounds, enzymes, active pharmaceutical compounds, substances for foodstuffs and food supplements, feedstuffs and feed supplements and cosmetics.

**[0126]** The substances here can be identical or different from one another that is to say with reference to the (chemical) components per se which make up the material, or else in the concentration of identical components in a material to be investigated.

#### Automated production

**[0127]** The control program of the individual synthesis modules determines directly the establishment of the production parameters. Module in this context is taken to mean a synthesis unit which is independent in principle, for example a liquid or solids metering unit or a filtration and washing unit. In the case of optimum system integration, the modules are controlled by a shared computer-aided control program or process control system and the individual sequences of the modules are controlled by suitable interaction of the sensors and actuators. The control system preferably causes a synchronization of sensors and actuators on a real-time basis. The production parameters and the definition or characteristics of the synthesis modules used are preferably stored as information in a database and in parallel thereto are converted into physical data (DA conversion) for controlling the synthesis modules. The information path to the database is generally accomplished via AD conversion and subsequent formatting of the data sequences.

#### Analysis

**[0128]** The characteristic properties of the substances ("1<sup>st</sup> and 2<sup>nd</sup> order properties") which can be obtained by various analytical methods, are also stored as information in the database. Depending on the analytical method used, the information is either of discrete or continuous nature. For example, the information as to whether a defined radiographic phase having particular performance characteristics is present or not is of discrete nature, whereas the report of the percentage content of this phase in the material is information of continuous or indiscreet nature.

### Automated testing

[0129] Automated testing initially requires definition of the test modules and establishing an initial set of the test parameters which unambiguously describes the test or the test sequence. Typical test parameters (for catalysts) comprise firstly test conditions such as concentration of the starting material stream, gas throughput/space velocity, temperature and pressure and secondly characteristic reactor data such as reactor type, heat conductance, dimensions and residence time characteristics, for example residence time and/or space velocity. In addition, depending on the type of substances in the substance library, electrical, thermal, mechanical, morphological, optical and magnetic test parameters can be determined, for example superconducting characteristics, critical current, critical magnetic field, conductivity, resistance, dielectric constants, strength and dielectric loss, polarization, piezoelectricity, expansion coefficient, thermal conductivity, vapour pressure; tension, anisotropy, adhesion, hardness, density, elasticity, porosity; crystallinity, microstructure, surface properties, crystal orientation; refractive index, absorption, birefringence, spectroscopic characteristics, emission, turbidity; magnetic resistance, coercivity, magnetic susceptibility; permeability, purity, complexing.

[0130] Ideally, the test parameters and the definition of the test modules are stored as information in a database and converted into physical data in parallel for controlling the test modules. In addition, there follows the simultaneous display and evaluation of the test results which are stored as information in the database. The test parameters and analytical data are generally of discrete (component/product present, yes or no) and continuous (concentration of the component) nature.

### Data capture

[0131] Data capture within the database provides that every substance of the substance library, in addition to a sample ID, is assigned all information from the automated synthesis, analysis and automated testing. For traceability, the information is preferably completed by specifying system

time, system location and the data. The unambiguous assignment of a sample and a result is made via a time signature and a reading unit, for example a barcode reader. Linking the automated production and testing and the analysis to a shared software environment ensures data security and data integrity.

#### Data analysis

**[0132]** Data analysis comprises the computer-aided evaluation of the acquired data using statistical methods. The purpose of data analysis is to look for correlations between the production parameters (characteristic properties) and the performance characteristics, to constrict the parameter spaces and thus to provide conditions for optimized substance libraries. The procedure in principle is shown in Figs. 5a and 5b. Fig. 5c shows diagrammatically the methodology of evolutionary algorithms for preparing optimized substance libraries. In this case the production parameters of those substances which have the best performance characteristics are mutated or combined in a systematic manner, so that production parameters which have been modified in this manner are prepared for the next generation of substances.

**[0133]** Step 1: The production (characteristic) parameters and test parameters (each line represents the testing of a defined substance within the substance library under defined conditions) are established within a defined parameter space.

**[0134]** Step 2: Depending on the production parameters (characteristic parameters), for each substance there is, under defined test conditions, characterized by a complete set of test parameters, a concrete test result (for heterogeneous catalysis this is generally conversion rate and selectivity as a function of T, p, feed gas composition and the type of the reactor).

**[0135]** Step 3: By using statistical methods, for example classic statistical methods, regression methods, linear or non-linear regression, data mining methods, neuronal networks or evolutionary methods, for example stepwise and multiple regression, factorial regression, polynomial regression, response surface regression, principal component analysis, partial least squares,

evolutionary algorithms, genetic function algorithm, the production and test parameters of the substances and if appropriate their characteristic properties are tested with respect to their effect on their performance characteristics. By this means the parameter space can be constricted and the search for optima of the target parameters can be restricted to the high-influence parameters. The choice of the preferred mathematical methods and algorithms is open in principle.

[0136] Step 4: By targeted experimenting, trends in selectivity analysis are verified.

[0137] Step 5: An optimized substance library is prepared and the cycle is run through again.

[0138] In addition, the present invention relates to a computer program having program code means for carrying out the inventive method, a data carrier containing precisely this computer program and a computer program of the type which can be carried out using the apparatus of the invention, in order to achieve implementation of the inventive method.

[0139] The present invention will now be further explained with reference to an example.

#### Example

[0140] The mode of functioning of the inventive method will be explained below with reference to an example: The objective was to find the optimum production and test parameters for a complex heterogeneously catalysed reaction. The reaction selected was selective NO<sub>x</sub> reduction by hydrocarbons at Cu-containing catalysts. The catalysts were produced by automated parallel ion exchange of  $\alpha$ -Al<sub>2</sub>O<sub>3</sub> and Al<sub>2</sub>O<sub>3</sub>/SiO<sub>2</sub> (zeolites) in Cu acetate. In this process 5 grams of  $\alpha$ -Al<sub>2</sub>O<sub>3</sub> or zeolite were charged with 0.5 l of Cu acetate solution and stirred for a further 5 hours at room temperature.

[0141] The catalytic reaction took place in a 48-tube tube-bundle reactor which was operated in a fully automatic synthesis gas test plant. The reaction took place at a space velocity of 50,000 h<sup>-1</sup> based on the active component. The hydrocarbon (HC) used was propene. The target parameter was NO<sub>x</sub>

conversion rate. The following production parameters and test parameters were varied within the specified limits:

Parameter	Parameter type	Parameter range
pH during ion exchange	Production parameter	0.68–11
CuAc <sub>2</sub> concentration during ion exchange	Production parameter	0.034–0.565 mol/l
Drying temperature	Production parameter	50–200 °C
Calcination temperature	Production temperature	200–700 °C
HC concentration	Test parameter	4.4–2395.6 ppm
Temperature	Test parameter	134.3–665.7 °C
Oxygen concentration	Test parameter	1.01–8.98% by volume

The support materials used were the following zeolite/  $\text{Al}_2\text{O}_3$ :

H-ZSM-5 (module ~ 15)	("Block 1")
H-ZSM-5 (module ~ 30)	("Block 2")
H-ZSM-5 (module ~ 140)	("Block 5")
NH4-Y (module ~ 6)	("Block 3")
$\alpha$ - $\text{Al}_2\text{O}_3$	("Block 4")

**[0142]** The orthogonal design (step 1) of the initial production and test library and the associated NOx conversion rates ("data acquisition", step 2) is shown in Table 2. The design provides the analysis of 7 parameters which are organized in five blocks (one block in each case represents the use of a zeolite or alumina) and in total 96 catalytic tests.

**[0143]** Using a second order regression model (step 3), the effects which the production and test parameters have on the NOx conversion rate are evaluated. Figure 6 shows a simplified presentation of the effects in the form of a Pareto diagram. Effects which prove to be statistically relevant are (with decreasing order of effect) the test temperature, the pH during ion exchange, the HC concentration in the test gas and the  $\text{CuAc}_2$  in the exchange solution. In addition, all zeolites/aluminas used differed significantly. Catalysts based on Cu-exchanged Y-zeolite or  $\alpha$ -alumina show only low NOx activity. The parameters  $\text{O}_2$  concentration, drying temperature and calcination temperature were of virtually no importance. These were kept constant in the subsequent experiments as follows:

$\text{O}_2$  concentration: 5% by volume

Drying temperature: 140°C

Calcination temperature: 500°C

[0144] The subsequent 1<sup>st</sup> optimized substance library, which comprises in total 48 experiments, takes into account only those parameters which have proved to be significant in the initial substance library and is restricted to the two most active zeolite materials Cu/H-ZSM-5 (M ~ 30) ("block 2") and Cu/H-ZSM-5 (M ~ 140) ("block 5"). The design of the substance library and the associated NOx conversion rates are shown in Table 3. The decrease in the width of variation for the parameters test temperature (from 250 to 550°C) and pH (from 4 to 8) is notable. The corresponding Pareto diagram in Fig. 7 identifies the parameters test temperature, HC concentration and pH as having particularly high effects. In addition, the data show the higher NOx activity of Cu/H-ZSM-5 (M ~ 30) ("block 2") compared with Cu/H-ZSM-5 (M ~ 140) ("block 5"), so that the last substance library design only takes into account the last-mentioned parameters (test temperature, HC concentration and pH) and Cu/H-ZSM-5 (M ~ 30) ("block 2") (Table 4). This is composed of a further 16 experiments (25 to 40), together with 24 experiments (1 to 24) which were carried out in the preceding substance library on Cu/H-ZSM-5 (M ~ 30) ("block 2"). The CuAc<sub>2</sub> concentration was kept constant for the new 16 experiments at 0.3 mol/l.

[0145] Using a simplex algorithm, the NOx conversion rate maximum was found at the following parameter setting: T = 418°C, HC concentration = 1698 ppm, pH for the Cu ion exchange: 6.1.

[0146] Moreover, the mathematical models mentioned above under "step 3" are fundamentally suitable for analyzing a substance library:

[0147] The discovery of extreme values within a parameter space can be carried out using the customary methods, for example the simplex method, the method of Powell, Quasi-Newton methods, or the method of "simulated annealing".

**Table 1****1. Production parameters of the most important preparation methods for catalysts**

<u>1. Precipitation/coprecipitation</u>	<u>2. (Metal) adsorption to support</u>	<u>3. Pore filling/incipient wetness</u>
- Precursor	- Precursor (interaction with the support)	- Precursor
- Support characteristics	- Support characteristics	- Support characteristics
- Temperature (temperature program etc.)	- Temperature (temperature program etc.)	- Agitation/mixing
- Precipitation reagents ( $\text{C}_2\text{O}_4^{2-}$ , $\text{CO}_3^{2-}$ , OH-urea etc.)	- Injection rate, agitation	- Concentrations
- Injection rate, agitation	- pH	- Viscosity of the precursor solution
- pH	- Concentrations	- Pressure
- Concentration	- Viscosity of the precursor solution	- Drying (static/flowing gas atmosphere)
- Ageing of the suspensions (time, temperature)	- Washing (pH, salts, temperature, volume, agitation)	- Calcination (reducing/oxidizing atmosphere, temperature rise, maximum temperature)
- Washing (pH, salts, temperature, volume, agitation)	- Drying (static/flowing gas atmosphere)	- Calcination (reducing/oxidizing atmosphere, temperature rise, maximum temperature)
- Drying (static/flowing gas atmosphere)	- Calcination (reducing/oxidizing atmosphere, temperature rise, maximum temperature)	- Calcination (reducing/oxidizing atmosphere, temperature rise, maximum temperature)
- Calcination (reducing/oxidizing atmosphere, temperature rise, maximum temperature)		

Table 2

2. Production parameters, test parameters and measured NO<sub>x</sub> conversion rate of the initial library, test 1 - 48



Table 3

3. Production parameters, test parameters and measured NOx conversion rate of the **1<sup>st</sup> optimised library**, parameter space was constricted and the variable range reduced

BLOCK	T_TEST	HC_CCNC	PH	CU_CONC
1	225,000	750,000	5,000	200
2	225,000	750,000	5,000	400
3	225,000	750,000	7,000	200
4	225,000	750,000	7,000	400
5	225,000	1650,000	5,000	200
6	225,000	1650,000	5,000	400
7	225,000	1650,000	7,000	200
8	225,000	1650,000	7,000	400
9	475,000	750,000	5,000	200
10	475,000	750,000	5,000	400
11	475,000	750,000	7,000	200
12	475,000	750,000	7,000	400
13	475,000	1650,000	5,000	200
14	475,000	1650,000	5,000	400
15	475,000	1650,000	7,000	200
16	475,000	1650,000	7,000	400
17	250,000	1200,000	6,000	200
18	550,000	1200,000	6,000	200
19	2	300,000	6,000	100
20	2	400,000	6,000	200
21	2	400,000	6,000	400
22	2	400,000	6,000	100
23	2	400,000	6,000	200
24	2	400,000	6,000	400
25	5	225,000	750,000	4,000
26	5	225,000	750,000	5,000
27	5	225,000	750,000	7,000
28	5	225,000	750,000	7,000
29	5	225,000	1650,000	5,000
30	5	225,000	1650,000	5,000
31	5	225,000	1650,000	7,000
32	5	225,000	1650,000	7,000
33	5	475,000	750,000	5,000
34	5	475,000	750,000	5,000
35	5	475,000	750,000	7,000
36	5	475,000	750,000	7,000
37	5	475,000	1650,000	5,000
38	5	475,000	1650,000	5,000
39	5	475,000	1650,000	7,000
40	5	475,000	1650,000	7,000
41	5	210,000	1200,000	6,000
42	5	150,000	1200,000	6,000
43	5	400,000	100,000	100
44	5	400,000	2100,000	100
45	5	400,000	1200,000	100
46	5	400,000	1200,000	100
47	5	400,000	1200,000	100
48	5	400,000	1200,000	100

**Table 4**

4. Production parameters, test parameters and measured NOx conversion rate of the 2<sup>nd</sup> optimized library, parameter space was further constricted and the variable range reduced, optimum (saddle point) determined via the simplex method

T <sub>TEST</sub>	HC CONC	pH	u <sub>NOX</sub>
1	325,000	5,000	17,000
2	325,000	5,000	21,000
3	315,000	5,000	13,000
4	315,000	5,000	17,000
5	315,000	5,000	26,000
6	315,000	5,000	31,000
7	313,000	5,000	19,000
8	312,000	5,000	24,000
9	312,000	5,000	41,000
10	310,000	5,000	45,000
11	310,000	5,000	41,000
12	310,000	5,000	18,000
13	310,000	5,000	51,000
14	310,000	5,000	52,000
15	310,000	5,000	48,000
16	310,000	5,000	49,000
17	250,000	1200,000	6,000
18	530,000	1200,000	6,000
19	490,000	200,000	6,000
20	490,000	2100,000	6,000
21	490,000	1200,000	4,000
22	490,000	1200,000	8,000
23	490,000	1200,000	6,000
24	490,000	1200,000	6,000
25	315,000	750,000	5,500
26	315,000	1650,000	5,500
27	315,000	750,000	6,000
28	315,000	1650,000	6,000
29	412,000	750,000	5,500
30	412,000	1650,000	5,500
31	412,000	750,000	6,000
32	412,000	1650,000	6,000
33	357,955	1200,000	6,150
34	441,043	1200,000	6,150
35	400,000	1200,000	5,057
36	400,000	1200,000	7,241
37	400,000	493,193	6,150
38	400,000	1956,107	6,150
39	400,000	1200,000	6,150
40	400,000	1200,000	6,150

Optimum:

T: 418°C

pH: 6.1

HC concentration: 1698 ppm

[0148] The present invention is by no int nt limited to the embodiment described heretofore, and modification may be made without departing from invention.

[0149] German priority application No. 10028875.8, filed June 10, 2000, including the specification, drawings, claims and abstract, is hereby incorporated by reference.